A Distance Measure for Attention Focusing and Anomaly Detection in Systems Monitoring

Richard J. Doyle

Artificial Intelligence Group
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109-8099
rdoyle@aig.jpl.nasa.gov

Content areas: Monitoring and diagnosis
automated reasoning, causal reasoning

January 22, 1994

A version of this paper also has been submitted to the *8th International Workshop on Qualitative Reasoning* (QR-94).

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

A Distance Measure for Attention Focusing and Anomaly Detection in Systems Monitoring

Content areas: Monitoring and diagnosis
automated reasoning, causal reasoning

January 22, 1994

Abstract

Any attempt to introduce automation into the monitoring of complex physical systems must start from a robust anomaly detection capability. This task is far from straightforward, for a single definition of what constitutes an anomaly is difficult to come by. In addition, to make the monitoring process efficient, and to avoid the potential for information overload on human operators, attention focusing must also be 'addressed. When an anomaly occurs, more often than not several sensors are affected, and the partially redundant information they provide can be confusing, particularly in a crisis situation where a response is needed quickly.

The focus of this paper is a new technique for attention focusing. The technique involves reasoning about the distance between two frequency distributions, and is used to detect both anomalous system parameters and "broken" causal dependencies. These two forms of information together isolate the locus of anomalous behavior in the system being monitored.

1 Introduction

Mission Operations personnel at NASA have the task of determining, from moment to moment, whether a space platform is exhibiting behavior which is in any way anomalous, which could disrupt the operation of the platform, and in the worst case, could represent a loss of ability to achieve mission goals. A traditional technique for assisting mission operators in space platform health analysis is the establishment of alarm thresholds for sensors, typically indexed by operating mode, which summarize which ranges of sensor values imply the existence of anomalies. Another established technique for anomaly detection is the comparison of predicted values from a simulation to actual values received in telemetry. However, experienced mission operators reason about more than alarm threshold crossings and discrepancies between predicted and actual to detect anomalies: they may ask whether a sensor is behaving differently than it has in the past, or whether a current behavior may lead to-the particular bane of operators—a rapidly developing alarm sequence.

Our approach to introducing automation into real-time systems monitoring is based on two observations: 1) mission operators employ multiple methods for recognizing anomalies, and 2) mission operators do not and should not interpret all sensor data all of the time. We seek an approach for determining from moment to moment which of the available sensor data is most informative about the presence of anomalies occurring within a system, The work reported here extends the anomaly detection capability in Doyle's SELMON monitoring system [3, 4] by adding an attention focusing capability.

Other model-based monitoring systems include Dvorak's MIMIC, which performs robust discrepancy detection for continuous dynamic systems [5], and DeCoste's DATMI, which infers system states from incomplete sensor data [2]. This work also complements other work within NASA on empirical and model-based methods for fault diagnosis of aerospace platforms [1,6, 7, 8].

2 Attention Focusing

A robust anomaly detection capability provides the core for monitoring, but only when this capability is combined with attention focusing does monitoring become both robust *and* efficient. Otherwise, the potential problems of information overload and too many false positives may defeat the utility of the monitoring system.

The attention focusing technique developed here uses two sources of information: historical data describing nominal system behavior, and causal information describing which pairs of sensors are constrained to be correlated, due to the presence of a dependency. The intuition is that the origin and extent of an anomaly can be determined if the misbehaving system parameters *and* the misbehaving causal, dependencies can be determined. Such information also supports reasoning to

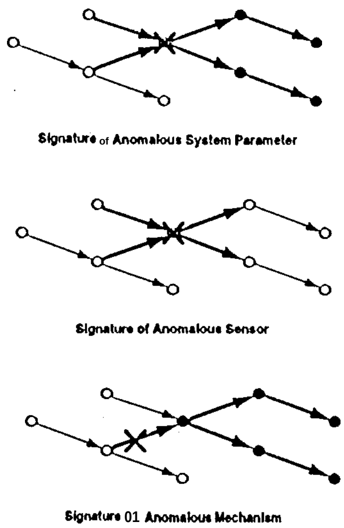

Figure 1: Anomalous System Parameters, Sensors and Mechanisms.

distinguish whether sensors, system parameters or mechanisms are misbehaving due to the fact that the signature of “broken” nodes and arcs in the causal graph are distinguishable. See Figure 1.

For example, the expected signature of an anomalous sensor includes the node of the sensor itself and the immediately adjacent arcs corresponding to the causal dependencies that the sensor participates in directly. The intuition is that the actual system is behaving normally so the locus of “brokenness” is isolated to the sensor and the set of adjacent causal dependencies which attempt to reconcile the bogus value reported by the sensor.

The expected signature of an anomalous system parameter also includes nodes and arcs which are downstream in the causal graph from the node corresponding to the system parameter. The intuition here is that the misbehavior, being in the actual system, will propagate.

The expected signature of an anomalous mechanism also includes arcs and nodes causally downstream from the arc corresponding to the mechanism. Once again, the intuition is that the misbehavior is in the system itself, and it will propagate. The way to distinguish this case from the anomalous system parameter case is to examine all input arcs (assuming there are more than one) to the most causally prior node in the “broken” subgraph.

2.1 Two Additional Measures

While SELMON runs, it computes incremental frequency distributions for all sensors being monitored. These frequency distributions can be' saved as a method for capturing behavior from any episode of interest. Of particular interest are historical distributions which correspond to nominal system behavior.

To identify an anomalous sensor, we apply a distance measure, defined below, to the frequency distribution which represents recent behavior to the historical frequency distribution representing nominal behavior, We call the measure simply *distance.* To identify a "broken" causal dependency, we first apply the same distance measure to the historical frequency distributions for the cause sensor and the effect sensor. This reference distance is a weak representation of the correlation that exists between the values of the two sensors due to the causal dependency. This reference distance is then compared to the distance between the frequency distributions based on recent data of the same cause sensor and effect sensor. The difference between the reference distance and the recent distance is the measure of the "brokenness" of the causal dependency. We call this measure *causal distance.*

2.2 Desired Properties of the Distance Measure

Define a distribution, D as the vector d_i such that

$$\forall i, 0 \leq d_i \leq 1$$

and

$$\sum_{i=0}^{n-1} d_i = 1$$

For a sensor S, we assume that the range of values for the sensor has been partitioned into n contiguous subranges which exhaust this range. We construct a frequency distribution as a vector D_S of length n, where the value of d_i *is* the frequency with which S has displayed a value in the ith subrange.

If our aim was only to compare different frequency distributions of the same sensor, we could use a distance measure which required the number of partitions, or bins in the two distributions to be equal, and the range of values covered by the distributions to be the same, However, since our aim is to be able to compare the frequency distributions of different sensors, these conditions must be relaxed.

Before defining the other desired properties of the distance measure, we define two special types of frequency distribution. Let F be the random, or flat distribution where $\forall i$, $d_i = \frac{1}{n}$. Let S_i be the set of "spike" distributions where $d_i = 1$ and $\forall j \neq i, d_j = 0$.

We seek a distance measure for frequency distributions with the following properties:

Distance
$\forall D_1 D_2, \Delta(D_1, D_2) \geq 0$
This property merely defines the measure as a distance measure.

Identity
$\forall D, \Delta(D, D) = 0$

Symmetry
$\forall D_1 D_2, \Delta(D_1, D_2) = \Delta(D_2, D_1)$
We do not wish to emphasize whether we are comparing recent data to historical data or vice versa, or cause data to effect data or vice versa.

Distinctness
$\forall D_1, D_2, \textit{ if } D_1 \neq D_2, \textit{then} \Delta(D_1, D_2) > 0$
The distance measure should distinguish distinct frequency distributions.

Spike Distinctness
$\forall i \neq j, \Delta(S_i, S_j) > 0$
We wish the set of S_i to be distinguishable.

Spike Ordering
$\forall i, \Delta(S_i, S_{i+1}) < \Delta(S_i, S_{i+2})$
The distance measure should preserve the fact that there is an ordering on the bins.

Spike Equidistance
$\forall i \neq j, \Delta(S_i, S_{i+1}) = \Delta(S_j, S_{j+1})$
There should be no difference in weighting of the spike distributions.

Spike/Flat Equidistance
$\forall i \neq j, \Delta(S_i, F) = \Delta(S_j, F)$
The difference between any spike distribution and the flat distribution is to be the same.

Extrema $\forall D_1 D_2 \forall i, \Delta(D_1, D_2) \leq \Delta(S_i, F)$
Any spike distribution and the flat distribution are to be considered the most different. All other distributions fall in between.

2.3 The Distance Measure

The distance measure is computed by projecting the two distributions into the two-dimensional space $[f, s]$ in polar coordinates and taking the euclidian distance between the projections,

Define the "flatness" component $f(D)$ of a distribution as follows:

$$\sum_{i=0}^{n-1} \frac{1}{2} \left| \frac{1}{n} - d_i \right|$$

This is simply the sum of the bin-by-bin differences between the given distribution and F. Note that O $\leq f(D) \leq$ 1. Also, $f(S_i) \rightarrow 1$ as n $\rightarrow \infty$.

Define the "spikeness" component s(D) of a distribution as:

$$\sum_{i=0}^{n-1} \phi \frac{i}{n-1} d_i$$

This is simply the centroid value calculation for the distribution. The weighting factor ϕ will be explained in a moment, Once again, O $\leq$ s(D) $\leq$ 1.

Now take [f,s] to be polar coordinates $[r, \theta]$. This maps F to the origin and the S_i to points along an arc on the unit circle., See Figure 2.

By inspection, the *Spike Distinctness, Spike Ordering* and *Spike/Flat Equidistance* properties are satisfied. The *Spike Equidistance* property is "satisfied because there is no unequal weighting applied in the centroid calculation, The *Distance, Identity* and *Symmetry* properties follow from taking the euclidian distance between the projections of the distributions. The *Extrema* property is satisfied by taking $\phi = \frac{\pi}{3}$. This choice of ϕ guarantees that $\Delta(S_0, S_{n-1}) = \Delta(F,$ SO) = $\Delta(F, S_{n-1})$ = 1 and all other distances in the region which is the range of A are by inspection ≤ 1.

The *Distinctness* property is not satisfied by the function A(D_1, D_2). This is not surprising because the multi-dimensional space arising from the number of bins in a distribution is collapsed to a two-dimensional space $[f, s]$. (Consider any two distributions D_1, D_2 with the same even number of bins such that the frequencies in the first $\frac{n}{2}$ bins and the frequencies in the second $\frac{n}{2}$ bins both sum to 0.5 in both D_1 and D_2. These two frequency sets within each distribution may be exchanged and/or permuted without violating A(D_1, D_2) = O). Thoughts on how to address this limitation appear below.

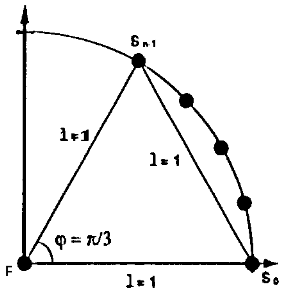

Figure 2: The function $\Delta(D_1, D_2)$.

Insensitivity to the number of bins in the two distributions and the range of values encoded in the distributions is provided by the $[f, s]$ projection function, which abstracts away from these properties of the distributions.

We may note in passing that the distance measure described here maybe easily modified to apply to continuous distributions, when theoretical models of the behavior of a system are available, The centroid calculation of the s component is easily accomplished, and the f component involves merely the integral of a difference, which may be accomplished numerically if necessary.

2.4 Results

In this section, we report on the results of applying the distribution distance measure to the task of focusing attention in monitoring. The distribution distance measure is used. to identify misbehaving nodes *(distance)* and arcs *(causal distance)* in the causal graph of the system being monitored, or equivalently, detect and isolate the extent of anomalies in the system being monitored,

2.4.1 A Space Shuttle Propulsion Subsystem

Figure 3 shows a causal graph for a portion of the Forward Reactive Control System (FRCS) of the Space Shuttle. A full causal graph for the Reactive Control System, comprising the Forward, Left and Right RCS, was developed with the domain expert.

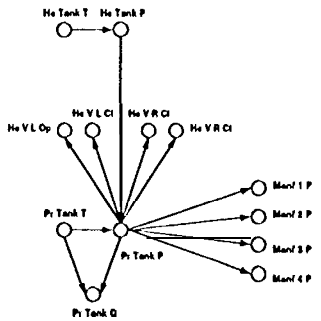

Figure 3: Causal Graph for the Forward Reactive Control System (FRCS) of the Space Shuttle.

2.4.2 **Attention Focusing Examples**

SELMON was run on seven episodes describing nominal behavior of the FRCS. The frequency distributions collected during these runs were merged. Reference distances were computed for sensors participating in causal dependencies.

SELMON was then run on 13 different fault episodes, representing faults such as leaks, sensor failures and regulator failures. Two of these episodes will be examined here; results were similar for all episodes, In each fault episode, and for each sensor, the distribution distance measure was applied to the incremental frequency distribution collected during the episode and the historical frequency distribution from the merged nominal episodes. These distances were a measure of the "brokenness" of nodes in the causal graph; i.e., instantiation of the *distance* measure.

New distances were computed between the distributions corresponding to sensors participating in causal dependencies. The differences between the new distances and the reference distances for the dependencies were a measure of the "brokenness" of arcs in the causal graph; i.e., instantiation of the *causal distance* measure.

The first episode involves a leak affecting the first and second manifolds (jets) on the oxidizer side of the FRCS. The pressures at these two manifolds drop to vapor pressure. The dependency between these pressures and the pressure in the propellant tank is severed because the valve between the propellant tank and the manifolds is closed. Thus there are two anomalous system parameters (the manifold pressures) and two anomalous mechanisms (the agreement between the

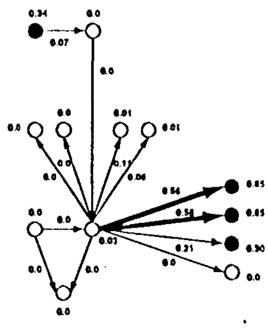

Figure 4: A leak fault,

propellant and manifold pressures when the valve is open).

The *distance* and *causal distance* measures computed for nodes and arcs in the FRCS causal graph reflect this faulty behavior. See Figure 4. (To visualize how the distribution distance measure circumscribes the extent of anomalies, the coloring of nodes and the width of arcs in the figure are correlated with the magnitudes of the associated distance and *causal distance* scores). An explanation for the apparent helium tank temperature anomaly is not available. However, we note that this behavior was present in the data for all six leak episodes.

The second episode involves an overpressurization of the propellant tank duc to a regulator failure. Onboard software automatically attempts to close the valves which isolate the helium tank from the propellant tank. One of the valves sticks and remains open.

The *distance* and *causal distance* measures isolate both the' misbehaving system parameters (propellant pressure and valve status indicators) and the altered relationships between the helium and propellant tank pressures and between the propellant tank pressure and the valve status indicators. Overpressurization of the propellant tank also alters the usual relation between propellant tank pressure and manifold pressures. See Figure 5.

3 Discussion

The *distance* and *causal distance* measures based on the distribution distance measure combine two concepts: 1) empirical data alone can bc an effective model of behavior, and 2) the existence of a causal dependency between two

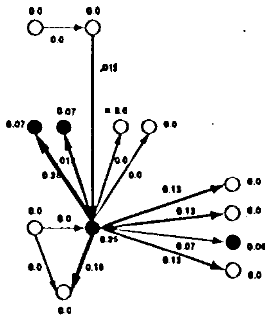

Figure 5: A pressure regulator fault,

parameters implies that their values are somehow correlated. The *causal distance* measure constructs a model of the correlation between two causally related parameters, capturing the general notion of constraint in an admittedly abstract manner. Nonetheless, these models of constraint arising. from causality provide surprising discriminatory power for determining which causal dependencies (and corresponding system mechanisms) are misbehaving. (In the *distance* measure for detecting misbehaving system parameters, wc are simply using the degenerate constraint of expected equality between historical and recent behavior.)

The approach described in this paper has usability advantages over other forms of model-based reasoning. The overhead involved in constructing the causal and. behavioral model of the system is minimal. The behavioral model is derived directly from actual data; no offline modeling is required. The causal model is of the simplest form, describing only the existence of dependencies. For the Shuttle RCS, a 198-node, 196-arc causal graph was constructed in a single one and one half hour session between the author and the domain expert,

3.1 Anomaly Characterization

Most model-based reasoning work has focused on diagnosis, treating monitoring as a "front-end", with discrepancy detection usually chosen as the monitoring technique. Our work suggests modifications to this view.

Monitoring is a complex, subtle and important task in its own right. The most sophisticated diagnosis engine is of limited utility if it is unreliably invoked by a weak anomaly detection module.

The monitoring/diagnosis distinction actually defines two poles of a continuum. At one end is anomaly detection. The goal of anomaly detection is simply to determine if an anomaly exists. General models of what constitutes an anomaly are utilized, with limited reference to explicit behavior models. Reasoning is local rather than global.

Next in the continuum is anomaly characterization. The goal here is to describe the extent of anomalous behavior. Again, the use of explicit behavior models is limited, but reasoning now encompasses a global view of the system. The anomaly detection capability of SELMON and the attention focusing capability which is the subject of this paper correspond to anomaly detection and anomaly characterization as defined here.

Next comes fault isolation. Reasoning now is refined from anomaly extent to anomaly source. Explicit behvaior models may be used, but not explicit fault models.

Finally comes full-fledged fault diagnosis, which includes an explanation of how the proposed fault produced the anomalous behavior. Explicit fault models may be referenced to verify hypotheses.

In actual real-time monitoring practice, operators perform anomaly detection and characterization routinely, and fault isolation when enough information is available to support their reasoning. Fault diagnosis is typically done off-line.

4 Future Work

Several issues need to be examined to continue the evaluation of the attention focusing technique based on the distribution distance measure and its utility in monitoring.

We need to understand the sensitivity of the technique to how sensor value ranges are partitioned. Clearly the discriminatory power of the distribution distance measure is related to the resolution provided by the number of bins and the bin boundaries. The results reported here are encouraging for the number of FRCS sensor bins were in many cases as low as three and in no cases more than eight.

We need to understand the suitability of the technique for systems which have many modes or configurations. We would expect that the discriminatory power of the technique would be compromised if the distributions describing behaviors from different modes were merged, Thus the technique requires that historical data representing nominal behavior is separable for each mode. If there are many modes, at the very least there is a data management task. A capability for tracking mode transitions is also required. An unsupervised learning system which can enumerate system modes from historical data and enable automated classification would solve this problem nicely.

We need to understand consequences of the *Distinctness* property not being satisfied by the distribution distance measure. Some distinct distributions arc not being distinguished; of more. relevant concern is whether or not distributions we wish to distinguish arc in fact being distinguished, The judicial introduction of additional components (e.g., the number of local maxima in a frequency distribution) to the distribution projection space $[f, s]$ may bc required to enhance discriminability.

The discriminatory power of the *causal distance* measure might be enhanced by retaining the flatness/spikeness distinction. For many linear functions, different input distributions may map to value-shifted but similarly shaped output distributions. In other words, the spikeness component may vary while the flatness component may be relatively invariant. It may bc possible to distinguish the case where misbehavior is the result of bogus values being propagated through still correctly functioning mechanisms.

It should be possible to describe the temporal (along with the causal/spatial) extent of anomalies by incrementally comparing recent sensor frequency distributions calculated from a "moving. window" of constant length with static reference frequency distributions.

5 Summary

We have described the properties and performance of a distance measure used to identify misbehavior at sensor locations and across mechanisms in a system being monitored. The technique enables the locus of an anomaly to bc determined. This attention focusing capability is combined with a previously reported anomaly detection capability in a robust, efficient and informative monitoring system, which is being applied in mission operations at NASA.

References

[1] K. H. Abbott, "Robust Operative Diagnosis as Problem Solving in a Hypothesis Space," *7th National Conference on Artificial Intelligence*, St. Paul, Minnesota, August 1988.

[2] D. DeCoste, "Dynamic Across-Time Measurement Interpretation," *8th National Conference on Artificial Intelligence*, Boston, Massachusetts, August 1990.

[3] R. J. Doyle, L. Charest, Jr., N. Rouquette, J. Wyatt, and C. Robertson, "Causal Modeling and Event-driven Simulation for Monitoring of Continuous Systems," *Computers in Aerospace 9*, American Institute of Aeronautics and Astronautics, San Diego, California, October 1993.

[4] R. J. Doyle, S. A. Chien, U. M. Fayyad, and E. J. Wyatt, "Focused Real-time Systems Monitoring Based on Multiple Anomaly Models," *7th International Workshop on Qualitative Reasoning*, Eastsound, Washington, May 1993.

[5] D. L. Dvorak and B. J. Kuipers, "Model-Based Monitoring of Dynamic Systems," *11th International Conference on Artificial Intelligence*, Detroit, Michigan, August 1989.

[6] T. Hill, W. Morris, and C. Robertson, "Implementing a Real-time Reasoning System for Robust Diagnosis' *6th Annual Workshop on Space Operations Applications and Research*, Houston, Texas, August 1992.

[7] J. Muratore, T. Heindel, T. Murphy, A. Rasmussen, and R. McFarland, "Space Shuttle Telemetry Monitoring by Expert Systems in Mission Control," in *Innovative Applications of Artificial Intelligence*, H. Schorr and A. Rappaport (eds.), AAAI Press, 1989.

[8] E. A. Scarl, J. R. Jamieson, and E. New, "Deriving Fault Location and Control from a Functional Model," *3rd IEEE Symposium on Intelligent Control*, Arlington, Virginia, 1988.